## REMARKS

In view of the following discussion, the Applicants submit that none of the claims now pending in the application is directed to non-statutory subject matter under 35 U.S.C. §101, unsupported under the provisions of 35 U.S.C. §112, anticipated under the provisions of 35 U.S.C. §102 or made obvious under the provisions of 35 U.S.C. §103. Thus, the Applicants believe that all of the presented claims are now in allowable form.

### I. ELECTION/RESTRICTION

The Applicants hereby affirm the election without traverse made by Kin-Wah Tong on February 14, 2005 to prosecute the invention of Group I, embodied in claims 1-31. Accordingly, claims 32-36 have been withdrawn without prejudice. The Applicants reserve the right to subsequently file one or more divisional applications in order to prosecute the invention recited in the non-elected group of claims.

### II. OBJECTION TO CLAIMS 18 AND 31

The Examiner has objected to claims 18 and 31 for informalities. In response, the Applicants have amended claims 18 and 31 to more clearly recite aspects of the present invention.

Specifically, the Applicants have amended both claims 18 and 31 to recite that "at least some of the distributions that indicate similarity are different", replacing "at least some of the distributions that indicate similarity differ".

Thus, the Applicants submit that the meanings of claims 18 and 31, as amended, are clear. Accordingly, the Applicants respectfully request that the objection to claims 18 and 31 be withdrawn.

### III. REJECTION OF CLAIM 29 UNDER 35 U.S.C. §101

The Examiner has rejected claim 29 under 35 U.S.C. §101 for being directed to non-statutory subject matter. In response, the Applicants have amended claim 29 to more clearly recite aspects of the present invention.

Specifically, the Applicants have amended claim 29 to recite "a computer readable medium containing an executable program for causing the processor to compare ...", replacing "software configured to cause the processor to compare ...". As noted by the Examiner, a program or software is statutory when embodied in a computer readable medium. MPEP 2106.

Thus, the Applicants submit that claim 29, as amended, fully satisfies the requirements of 35 U.S.C. §101. Accordingly, the Applicants respectfully request that the rejection to claim 29 under 35 U.S.C. §101 be withdrawn.

## IV. REJECTION OF CLAIM 28 UNDER 35 U.S.C. § 112

The Examiner has rejected claim 28 in the Office Action for allegedly being indefinite under 35 U.S.C. §112, second paragraph. In response, the Applicants have amended claim 28 in order to more clearly recite aspects of the present invention.

Specifically, the Applicants have amended claim 28 to recite a "machine readable medium carrying a model ...", replacing a "medium carrying a model". The Applicants submit that this amendment clarifies how the model is being carried, as requested by the Examiner.

Thus, the Applicants submit that claim 28 fully satisfies the requirements of 35 U.S.C. §112. Accordingly, the Applicants respectfully request that the rejection to claim 28 under 35 U.S.C. §112 be withdrawn.

## V. REJECTION OF CLAIMS 1-5, 7, 12, 15-23, 25 AND 27-31 UNDER 35 U.S.C. § 102

The Examiner rejected claims 1-5, 7, 12, 15-23, 25 and 27-31 under 35 U.S.C. §102(e) as being unpatentable over the Sundaresan patent (United States Patent No. 6,606,620, issued on August 12, 2003, hereinafter "Sundaresan"). The Applicants respectfully traverse the rejection.

Sundaresan teaches a method for classifying semi-structured documents such as World Wide Web pages. For example, the method may be implemented in a manner similar to an Internet search engine, in which the method retrieves appropriate web pages in response to user search queries. Specifically, for each document (*e.g.*, web

page), the method creates a hierarchical model of vectors and sub-vectors that encode the document's structural and textual information. At each vector or sub-vector, the frequency of the associated terms (*e.g.*, words appearing in the document) is calculated. The document is then compared to known classes of documents and is classified based the document's term frequency and term distribution characteristics. Classes of documents comprise normalized sums of individual vectors for a plurality of documents in that particular class.

The Examiner's attention is directed to the fact that Sundaresan fails to disclose or suggest the novel invention of specifying a model comprising a distribution of vectors, where at least some of the distributions indicate dissimilarity, similarity or match at particular positions of structured data objects, as claimed in Applicants' independent claims 1, 28, 29 and 30. Specifically, Applicants' claims 1, 28, 29 and 30 positively recite:

> 1. A method comprising:
> specifying a model that (i) represents a set of structured data objects that include elements at particular positions, and (ii) comprises distributions of vectors, each distribution corresponding to particular positions in the respective structured data objects, each of the vectors comprising values for the particular positions, at least some distributions indicating dissimilarity at particular positions of the structured data objects; and
> comparing a given set of structured data objects to the model to determine a likelihood that the given set is represented by the model. (Emphasis added)
>
> 28. A machine readable medium carrying a model capable of enabling a machine to perform comparisons of a set of structured data objects to the model, the model comprising distributions of vectors, each distribution corresponding to particular positions in the respective structured data objects such that each of the vectors comprises values for the particular positions, wherein at least some distributions indicating dissimilarity at particular positions of the structured data objects. (Emphasis added)
>
> 29. An apparatus comprising a processor and storage, the storage comprising (i) a model that represents a set of structured data objects, the model comprising distributions of vectors, each distribution corresponding to particular positions in the respective structured data objects such that each of the vectors comprises

values for the particular positions, wherein at least some distributions indicating dissimilarity at particular positions of the structured data objects, and (ii) a computer readable medium containing an executable program for causing the processor to compare a given set of structured data objects to the model. (Emphasis added)

30. A method comprising:

specifying a model that (i) represents a set of structured data objects that include elements at particular positions, and (ii) comprises distributions of vectors, each distribution corresponding to particular positions in the respective structured data objects such that each of the vectors comprises values for the particular positions, wherein at least some distributions indicate similarity between the structured data objects at particular positions and at least some others indicate matching to a reference structure data object at particular positions; and

comparing a given set of structured data objects to the model to determine a likelihood that the given set is represented by the model. (Emphasis added)

Applicants' invention is directed to a model for comparing structured data objects, *e.g.*, to determine the extent of similarity and to identify differences. Such comparison is useful, for example, to impute functions of the unidentified structured data object. However, the comparison of structured data objects is a complex process. For example, some structured data objects such as biopolymers (*e.g.*, nucleic acids, proteins, etc.) have complex sequences in which a given position in the sequence may include one of a number of potential elements (*e.g.*, nucleotides, amino acids, etc.). The properties of these potential elements are unique and varied, further complicating the comparison.

The present invention provides a method for classifying structured data objects by comparing the structured data objects to a model. The model is a reference that represents a set of structured data objects that include elements at particular positions in their respective sequences. Specifically, the model represents the relationships between these sequence positions. To this end, the model comprises a distribution of vectors, where each distribution corresponds to particular positions in the respective sequences and each vector comprises a value for the particular positions. At least some of the distributions indicate dissimilarity, similarity or a match between the set of

structured data objects at particular sequence positions. An unidentified structured data object (or set of structured data objects) may thus be compared to this model to determine the likelihood that the model represents the unidentified structured data object.

In contrast, Sundaresan teaches a model that simply calculates the frequencies with which certain terms appear throughout documents of a given type. Thus, Sundaresan fails to anticipate all of the limitations of the Applicants' invention.

Specifically, Sundaresan only teaches indicating a normalized sum or an aggregate number of appearances of a given term, which does not amount to indicating a match state (*e.g.*, dissimilarity, similarity or match) or substantive relationship between corresponding elements at specific positions in a set of structured data objects. The indication of a match state may help to impute or imply the function of a given structured data object, such as a biopolymer, based on a degree of similarity or dissimilarity or a match at specific positions to known structured data objects or models of known structured data objects. Sundaresan does not address the need to indicate a match state of two or more structured data objects at specific corresponding positions for use in a classification model. Sundaresan thus fails to teach or anticipate a method in which a model comprising a distribution of vectors is specified, wherein at least some of the distributions in the model indicating dissimilarity, similarity or match at particular positions of structured data objects, as positively claimed by the Applicants in claims 1, 28, 29 and 30. Therefore, the Applicants submit that, for at least the reasons stated above, independent claims 1, 28, 29 and 30 fully satisfy the requirements of 35 U.S.C. §102 and are patentable thereunder.

Dependent claims 2-5, 7, 12, 15-23, 25, 27 and 31 depend from claims 1, 28, 29 and 30, and recite additional features therefore. As such, and for at least the same reasons set forth above with respect to the rejection of independent claims 1, 28, 29 and 30, the Applicants submit that claims 2-5, 7, 12, 15-23, 25, 27 and 31 are not anticipated by the teachings of Sundaresan. Therefore, the Applicants submit that dependent claims 2-5, 7, 12, 15-23, 25, 27 and 31 also fully satisfy the requirements of 35 U.S.C. §102 and are patentable thereunder. Accordingly, the Applicants respectfully

request that the rejection to claims 1-5, 7, 12, 15-23, 25 and 27-31 under 35 U.S.C. §102 be withdrawn.

## VI. REJECTION OF CLAIMS 6, 8-9, 10-11, 13-14, 24 AND 26 UNDER 35 U.S.C. § 103

### 1. Claims 10-11, 13-14, 24 and 26

The Examiner rejected claims 10-11, 13-14, 24 and 26 under 35 U.S.C. §103(a) as being unpatentable over Sundaresan in view of the Krogh et al. article ("Hidden Markov Models In Computational Biology: Application to Protein Modeling", published on August 17, 1993, hereinafter "Krogh"). The Applicants respectfully traverse the rejection.

Sundaresan has been discussed above. Krogh teaches the application of Hidden Markov Models (HMMs) to the problems of statistical modeling, database searching and multiple sequence alignment of protein families and protein domains. The parameters of an HMM are estimated from a training set of unaligned sequences, and then used to search a database for other sequences that are members of a given family or that contain a given domain. In this way, the HMM distinguishes between members and non-members for given families.

The Examiner's attention is directed to the fact that Krogh, like Sundaresan, fails to disclose or suggest the novel invention of specifying a model comprising a distribution of vectors, where at least some of the distributions indicate dissimilarity, similarity or match at particular positions of structured data objects, as claimed in Applicants' independent claim 1, from which claims 10-11, 13-14, 24 and 26 depend. Applicants' claim 1 has been recited above.

As discussed above, Applicants' invention is directed to a model for comparing structured data objects such as biopolymers. Specifically, one embodiment of the Applicants' invention compares a given structured data object to a model that represents a set of structured data objects that include elements at particular positions in their respective sequences. Specifically, the model represents the relationships between these sequence positions. To this end, the model comprises a distribution of vectors, where each distribution corresponds to particular positions in the respective

sequences and each vector comprises a value for the particular positions. At least some of the distributions indicate dissimilarity, similarity or a match between the set of structured data objects at particular sequence positions.

As also discussed above, Sundaresan fails to teach or suggest a method in which a model comprising a distribution of vectors is specified, wherein at least some of the distributions in the model indicating dissimilarity, similarity or match at particular positions of structured data objects, as positively claimed by the Applicants in claim 1. Krogh fails to bridge this gap in the teachings of Sundaresan. Therefore, the Applicants submit that, for at least the reasons stated above, independent claim 1 fully satisfies the requirements of 35 U.S.C. §103 and is patentable thereunder.

Dependent claims 10-11, 13-14, 24 and 26 depend from claim 1, and recite additional features therefore. As such, and for at least the same reasons set forth above with respect to the rejection of independent claim 1, the Applicants submit that claims 10-11, 13-14, 24 and 26 are not made obvious by the teachings of Sundaresan in view of Krogh. Therefore, the Applicants submit that dependent claims 10-11, 13-14, 24 and 26 also fully satisfy the requirements of 35 U.S.C. §103 and are patentable thereunder. Accordingly, the Applicants respectfully request that the rejection to claims 10-11, 13-14, 24 and 26 under 35 U.S.C. §103 be withdrawn.

**2. Claims 6 and 8-9**

The Examiner rejected claims 6 and 8-9 under 35 U.S.C. §103(a) as being unpatentable over Sundaresan in view of the Mi/ke patent (United States Patent No. 5,787,414, issued on July 28, 1998, hereinafter "Mi/ke"). The Applicants respectfully traverse the rejection.

Sundaresan has been discussed above. Mi/ke teaches a method for retrieving data. Specifically, Mi/ke teaches a system that stores and retrieves primary data (*e.g.*, the target data to be stored/retrieved) using a retrieval key that is based on secondary information of the primary data (*e.g.*, environmental, structural or other information not directly related to the content of the primary data). The primary data is stored in a first memory, while the associated secondary information is stored in a corresponding

second memory. When a retrieval key is entered (*e.g.*, specifying a desired retrieval of the primary data from the first memory), the associated secondary information is selected from the second memory in accordance with the retrieval key, and the corresponding primary data is retrieved and output as a retrieval result.

The Examiner's attention is directed to the fact that Miike, like Sundaresan, fails to disclose or suggest the novel invention of specifying a model comprising a distribution of vectors, where at least some of the distributions indicate dissimilarity, similarity or match at particular positions of structured data objects, as claimed in Applicants' independent claim 1, from which claims 6 and 8-9 depend. Applicants' claim 1 has been recited above.

As discussed above, Sundaresan fails to teach or suggest a method in which a model comprising a distribution of vectors is specified, wherein at least some of the distributions in the model indicating dissimilarity, similarity or match at particular positions of structured data objects, as positively claimed by the Applicants in claim 1. Miike fails to bridge this gap in the teachings of Sundaresan. Therefore, the Applicants submit that, for at least the reasons stated above, independent claim 1 fully satisfies the requirements of 35 U.S.C. §103 and is patentable thereunder.

Dependent claims 6 and 8-9 depend from claim 1, and recite additional features therefore. As such, and for at least the same reasons set forth above with respect to the rejection of independent claim 1, the Applicants submit that claims 6 and 8-9 are not made obvious by the teachings of Sundaresan in view of Miike. Therefore, the Applicants submit that dependent claims 6 and 8-9 also fully satisfy the requirements of 35 U.S.C. §103 and are patentable thereunder. Accordingly, the Applicants respectfully request that the rejection to claims 6 and 8-9 under 35 U.S.C. §103 be withdrawn.

## VII. CONCLUSION

Thus, the Applicants submit that all of the presented claims now fully satisfy the requirements of 35 U.S.C. §101, §112, §102 and §103. Consequently, the Applicants believe that all of the presented claims are presently in condition for allowance. Accordingly, both reconsideration of this application and its swift passage to issue are

earnestly solicited.

If, however, the Examiner believes that there are any unresolved issues requiring the issuance of a final action in any of the claims now pending in the application, it is requested that the Examiner telephone Mr. Kin-Wah Tong, Esq. at (732) 530-9404 so that appropriate arrangements can be made for resolving such issues as expeditiously as possible.

Respectfully submitted,

May 23, 2005
Date

Kin-Wah Tong, Attorney
Reg. No. 39,400
(732) 530-9404

Moser, Patterson & Sheridan, LLP
595 Shrewsbury Avenue
Shrewsbury, New Jersey 07702